NASA/TM—2015-218478

Investigation of Spiral Bevel Gear Condition Indicator Validation Via AC–29–2C Combining Test Rig Damage Progression Data With Fielded Rotorcraft Data

Paula J. Dempsey
Glenn Research Center, Cleveland, Ohio

February 2015

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at *http://www.sti.nasa.gov*

- E-mail your question to *help@sti.nasa.gov*

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM—2015-218478

Investigation of Spiral Bevel Gear Condition Indicator Validation Via AC–29–2C Combining Test Rig Damage Progression Data With Fielded Rotorcraft Data

Paula J. Dempsey
Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Glenn Research Center
Cleveland, Ohio 44135

February 2015

This work was sponsored by the Fundamental Aeronautics Program at the NASA Glenn Research Center.

Level of Review: This material has been technically reviewed by technical management.

Available from

NASA STI Information Desk
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199

National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312

Available electronically at http://www.sti.nasa.gov

Investigation of Spiral Bevel Gear Condition Indicator Validation Via AC–29–2C Combining Test Rig Damage Progression Data With Fielded Rotorcraft Data

Paula J. Dempsey
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio 44135

Executive Summary

This report documents work performed under a NASA Space Act Agreement (SAA) with the Federal Aviation Administration (FAA) to support validation and demonstration of rotorcraft Health and Usage Monitoring Systems (HUMS) for maintenance credits via *FAA Advisory Circular (AC) 29–2C, Section MG–15, Airworthiness Approval of Rotorcraft (HUMS)* (Ref. 1). The overarching goal of this work was to determine methods to validate condition indicators in the lab that better represent their response to faults in the field.

The purpose of this work was to use existing in-service helicopter HUMS flight data from faulted spiral bevel gear sets as a "Case Study," to better understand differences in response between a helicopter and a component test rig. Spiral bevel gears, consisting of a pinion and gear, were chosen due to the availability of faulted spiral bevel gear condition indicator data from fielded helicopters and the availability of a test rig capable of performing damage progression tests on spiral bevel gears.

This is the final of three reports published on the results of this project. In the first report, results were presented on nineteen tests performed in the NASA Glenn Spiral Bevel Gear Fatigue Test Rig on spiral bevel gear sets designed to simulate helicopter fielded failures (Ref. 1). In the second report, fielded helicopter HUMS data from 40 helicopters were processed with the same techniques that were applied to spiral bevel rig test data (Ref. 2). Twenty of the 40 helicopters experienced damage to the spiral bevel gears, while the other 20 helicopters had no known anomalies within the time frame of the datasets. Refer to References 1 and 2 for results of the analyses on the test rig and helicopter HUMS data. In this report, results from the rig and helicopter data analysis will be compared for differences and similarities in condition indicator (CI) response. Observations and findings using sub-scale rig failure progression tests to validate helicopter gear condition indicators will be presented.

In the helicopter, gear health monitoring data was measured when damage occurred and after the gear sets were replaced at two helicopter regimes. For the helicopters or tails, data was taken in the flat pitch ground 101 percent rotor speed (FPG101) regime. For nine tails, data was also taken at 120 knots true airspeed (120KTA) regime. In the test rig, gear sets were tested until damage initiated and progressed while gear health monitoring data and operational parameters were measured and tooth damage progression documented. For the rig tests, the gear speed was maintained at 3500 rpm, a 1 hr run-in was performed at 4000 in-lb gear torque, than the torque was increased to 8000 in-lb.

The HUMS gear condition indicator data evaluated included Figure of Merit 4 (FM4), Root Mean Square (RMS) or Diagnostic Algorithm 1 (DA1), ± 3 Sideband Index (SI3) and ± 1 Sideband Index (SI1). These were selected based on their sensitivity in detecting contact fatigue damage modes from analytical, experimental and historical helicopter data. For this report, the helicopter dataset was reduced to fourteen tails and the test rig data set was reduced to eight tested gear sets. The damage modes compared were separated into three cases. For case one, both the gear and pinion showed signs of contact fatigue or

scuffing damage. For case two, only the pinion showed signs of contact fatigue damage or scuffing. Case three was limited to the gear tests when scuffing occurred immediately after the gear run-in.

Pearson Correlation Coefficients (r) were calculated to measure the strength and direction of the linear relationship between condition indicators and operational parameters for the test rig and helicopters. Details can be found in References 1 and 2. For both the test rig and helicopter, DA1/RMS values were sensitive to operational conditions. Both systems also showed correlations between right and left gear and pinion DA1/RMS values. SI1 and SI3 for the same component also had high r values. This response illustrates the importance of knowing how a specific CI is calculated per sensor response to minimize false alarms.

For ease of comparison, the test rig and helicopter CI data was separated into comparable failure modes. Three rig tests and nine helicopters with gear and pinion damage were compared. Two rig tests and five helicopters with only pinion damage were compared. For the pinion and gear damage case, DA1/RMS detection rate was better in the helicopter while FM4 performed better in the test rig and was detected by the gear CI, not the pinion, for both the helicopter and the test rig. SI1 performed the best at detecting gear and pinion damage for both the test rig and the helicopter. For the pinion only damage case, SI1 and SI3 performed better in the helicopter. The measured DA1/RMS, SI1 and SI3 values were significantly higher in the helicopter than the test rig. This was predicted by vibration transfer path measurements made to characterize each system that indicated lower amplitudes of frequency response functions measured on the test rig than the helicopter. The effect test rig design had on vibration response was also determined by these measurements.

Many factors can affect the measured vibration response within the system. A good gear condition indicator must be sensitive to specific failure modes and insensitive to environmental and operating conditions. The factors that can potentially affect CI response were discussed and a process flowchart was presented with the steps required for test rig condition indicator validation. The final step in this process was the quantification of CI performance metrics using damage detection rates and false alarm rates demonstrated for the test rig data. This process also showed SI1 performed the best at detecting spiral bevel gear set damage for both systems.

Results of this investigation highlighted the importance of understanding the complete monitored systems, for both the helicopter and test rig, before interpreting health monitoring data. Further work is required to better define these two systems that include better state awareness of the fielded systems, new sensing technologies, new experimental methods or models that quantify the effect of system design on CI response and new methods for setting thresholds that take into consideration the variance of each system.

Contents

1.0 Background

This report documents the results of an analysis on gear condition indicator data collected on helicopters and in a test rig when damage occurred to spiral bevel gear sets. Spiral bevel gears or gear sets consist of a pinion and a gear. Spiral bevel gears are used in a helicopter to modify rotational speed, torque and direction of the gas turbine power. Spiral bevel gears in helicopter transmissions transfer power between nonparallel intersecting shafts.

This analysis was performed as a collaborative effort with the Federal Aviation Administration (FAA) and the U.S. Army. The purpose of this work was to support validation and demonstration of rotorcraft Health and Usage Monitoring Systems (HUMS) for maintenance credits via *FAA Advisory Circular (AC) 29–2C, Section MG–15, Airworthiness Approval of Rotorcraft (HUMS)* (Ref. 3). Maintenance credits are modified inspection and removal criteria of components based on HUMS measured condition and actual usage. Maintenance credit validation includes providing evidence of damage detection algorithm effectiveness by using acceptance limits, trending techniques, tests and demonstration methods. For dynamic transmission components, these methods can include using data from naturally occurring aircraft faults and component fault testing in a test stand (Ref. 4). Due to time, cost and safety concerns, evidence via actual service on aircraft is typically replaced with rig tests, where a measurable and known component fault is checked against the algorithm and its thresholds.

The United Kingdom (UK) Civil Aviation Authority (CAA) also published a document to provide guidance using vibration health monitoring, defined as "data generated by processing vibration signals to detect incipient failures or degradation of mechanical integrity," for maintaining helicopter rotor and drive systems (Ref. 5). These vibration signals or signatures are referred to as "condition indicators" (CI) that develop when a fault occurs on a component and interacts with its operational environment. Within this CAA document, fault testing is also mentioned as a validation method to demonstrate algorithm damage detection effectiveness for specific faults.

The overarching goal of this work was to define condition indicator validation methods in the test rig that better represent their response to faults in the field. Response of a CI to a fault in a test rig may not always be representative of a CI response in a helicopter due to differences in both systems and their operational environments. For these situations, CI performance limitations must be defined to understand the risks in using a test rig validated CI on a helicopter.

In a previous analysis, CI performance was evaluated using fielded helicopter datasets recorded when damage occurred on spiral bevel gear (pinion/gear) teeth located in several helicopter nose gearboxes (Refs. 6 and 7). These nose gearboxes are mounted in the front of each engine to reduce speed and change the drive angle. Thousands of hours of CI data collected before and after spiral bevel gear replacement combined with tear down analyses (TDA) documenting the extent of damage were analyzed. Within this timeframe, NASA Glenn Research Center had an existing and available component test rig for testing spiral bevel gears, the Spiral Bevel Gear Fatigue Test Rig.

Using this existing in-service HUMS flight data from faulted spiral bevel gears, a plan was put in place to design and test comparable gear sets within the constraints of the test rig. Testing in the component test rig, as opposed to a full-scale system, were based on the cost and time it would take to design, develop and build testing capabilities combined with the time required to initiate and progress a defect in the actual helicopter component. The availability of fielded helicopter HUMS CI data, the Spiral Bevel Gear Fatigue Rig and the same HUMS installed on the helicopters for use in the rig made this a cost effective proposition within a reasonable timeframe. The requirements for testing spiral bevel gears in the Spiral Bevel Gear Fatigue Test Rig are outlined in Reference 8. Results of these tests can be found

in Reference 1. Fielded helicopter HUMS data from 40 helicopters, 20 with spiral bevel gear damage and 20 with no known anomalies, was also reprocessed with the same techniques applied to spiral bevel rig test data (Ref. 2).

Results from the analysis of the helicopter data (Ref. 2) and the rig data (Ref. 1) were combined for comparison and analysis in this report. Test rig damage progression data was compared to fielded rotorcraft data to assess the similarities and differences in gear condition indicator response between both systems. Observations and findings using subscale rig failure progression tests to validate helicopter gear condition indicators are then discussed within this report.

2.0 Objectives and Approach

The objective of the analysis was to support validation and demonstration of HUMS for maintenance credits via AC-29-2C (Ref. 3). The main focus was the use of rig tests as a validation method for HUMS drive train mechanical component condition indicator performance. CIs can perform differently when taken from a component test stand or even a full-scale transmission test stand and implemented on an aircraft. If CI data collected from a test rig is to be used to validate CI performance, it must be shown that its performance can be maintained when installed on the aircraft.

The presented approach compared in-service HUMS CI data from fielded helicopters with damaged spiral bevel gears to those tested in the NASA Spiral Bevel Gear Fatigue Test Rig. The tested spiral bevel gears were designed with comparable material and heat treatments as the helicopter gears. The damage in the test rig was not seeded, but allowed to initiated naturally and progress to failure modes comparable to those seen in the field.

The specific objectives were to compare HUMS CI performance between the helicopter and test rig in detecting gear surface pitting fatigue and other generated failure modes on spiral bevel gear teeth. A more general research objective was to gain a better understanding of CI validation using rig testing and to answer questions such as what can we scale, simulate and quantify with rig testing for validation of field performance. Current HUMS gear condition indicators were assessed. These include Figure of Merit 4 (FM4), Root Mean Square (RMS) or DA1, ± 1 Sideband Index (SI1) and ±3 Sideband Index (SI3) or SI. Comparisons were also made between the effect of flight regimes on helicopter CI performance and the effect of operational conditions on test rig CI performance.

3.0 Helicopter and Test Rig Spiral Bevel Gears

The helicopter component investigated was the spiral bevel gear set in the nose gearbox of the Apache helicopter (AH64). A photograph of the nose gearbox location in the helicopter is shown in Figure 1 (Ref. 8). Building a rig around an existing AH64 helicopter gear box was unfeasible for this project due to the cost and time to build a new rig, limited AH64 gears for test and the time required to initiate and progress fatigue failures in the gear sets. For these reasons, rig tests were performed in the Spiral Bevel Gear Fatigue Test Rig at NASA Glenn Research Center. A detailed description of this test facility is provided in References 9 and 10. The Spiral Bevel Gear Fatigue Test Rig is illustrated in Figure 2. A cross sectional view is also shown in Figure 2. The facility operates as a closed-loop torque regenerative system, where the drive motor only needs to supply enough power to overcome power losses within the system. The load is locked into the loop via a split shaft and a thrust piston that forces a floating helical gear axially into mesh. A 100 horsepower drive motor supplies the facility with rotation and loop losses.

Figure 1.—Location of spiral bevel gear set in helicopter

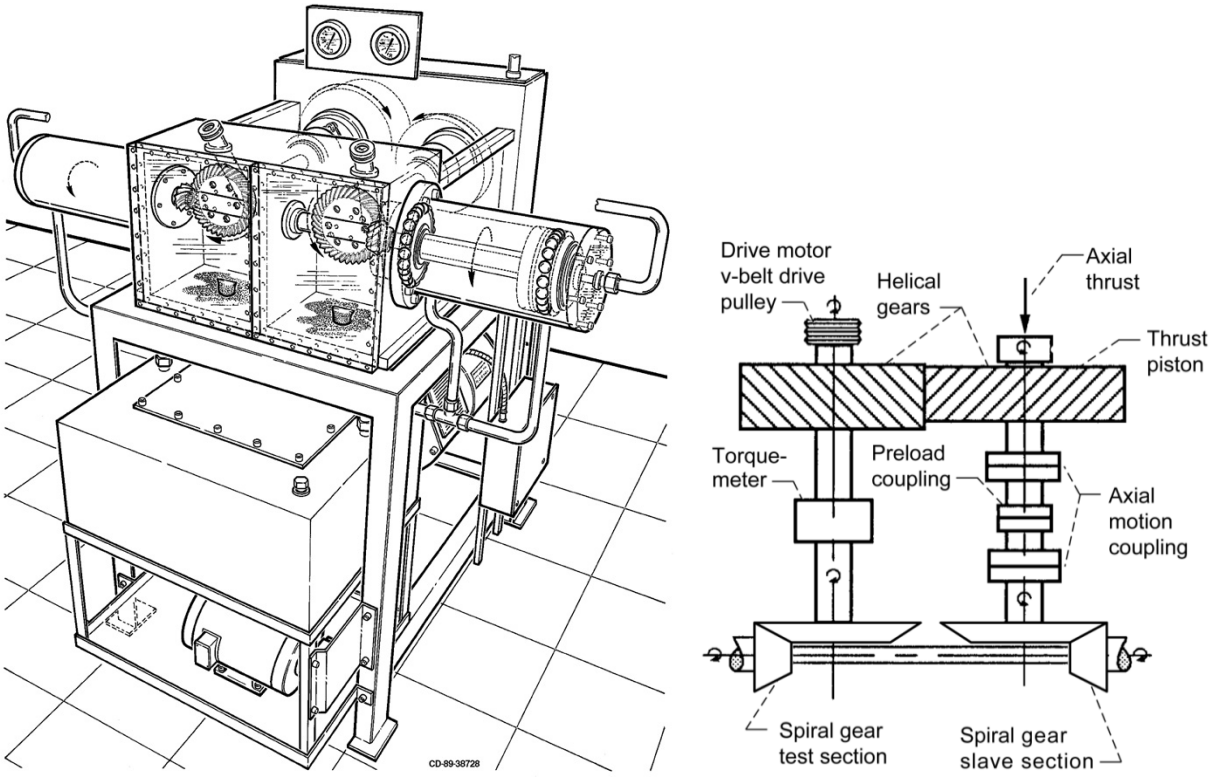

Figure 2.—Spiral Bevel Gear Fatigue Test Rig.

Two sets of spiral bevel gears are installed in the test rig and tested simultaneously. Facing the gearboxes, the left gear set is referenced as left and the right gear set is referenced as right within this report. Although the concave side of the pinion is always in contact with the convex side of the gear on either side, the left side pinion drives the gear in a normal speed reducer mode, while the right side gear drives the pinion.

The production AH64 gear design could not be directly adapted into the test rig due to the size of the gear set, the speed in which it runs and the test stand loading capacity. For this reason, the helicopter OEM was tasked to evaluate a gear set design that could be used in the test rig with close approximation/simulation to the helicopter gear set. The design met the following requirements:

1. The gear set design was constrained to the space available in the spiral bevel gear fatigue rig gearboxes within its speed and load limitations.
2. The gear sets were designed to fail in a manner comparable to the failure modes observed on the helicopters determined through review of the tear down analysis (TDA) and on-site inspection of several of the gearboxes.
3. The gear sets were designed to insure fatigue failures with limited overall test time.

A model was developed using AGMA software (Ref. 11) that identified the physics of failure based on the gear design and operating conditions. The program uses design information that includes speeds, torques, gear materials, heat treatments, lubricants, operating temperatures and reliability to determine contact stress indices and margins of safety. The final design was selected with the highest probability of failure in a pitting mode within 100 hr of operation at rig speed and load conditions, while avoiding bending fatigue failures. For the final test rig gear set design, the helicopter gear set material, gear ratio, pressure angle, spiral angle and contact stress were maintained, while the diametral pitch varied.

4.0 Instrumentation and Data Acquisition

Both the aircraft and test rig CI data was collected and processed with a helicopter HUMS, referred to as the Modern Signal Processing Unit (MSPU). The MSPU system is an on-board rotorcraft HUMS system that acquires, digitizes and processes the tachometer pulses and accelerometer data. The data is then downloaded to a ground station, referred to as Personal Computer Ground Based Station (PC-GBS), for further analysis. The ground station software provides the graphical user interface for the user to archive, process and analyze the on-board data.

Although the MSPU monitors all of the dynamic mechanical components in the helicopter, the focus for this analysis was limited to the components in the nose gearbox of the helicopter. The accelerometer frequency range was 0.5 to 5 KHz with a resonant frequency of 26 KHz. One magnetic tachometer was installed and used to measure shaft speed. Gear ratios were used to process the data at the correct speed for each component.

For the helicopter MSPU system, one accelerometer was installed on the housing of the right nose gearbox, and one was installed on the housing of the left nose gearbox. The accelerometers were mounted radially with respect to the gear. The magnetic tachometer measured speed at the main transmission. The gear ratio is then used to process the data at the correct speed for the gear.

For the test rig MSPU system, one accelerometer was installed on the housing of the right gearbox, and one was installed on the housing of the left gearbox. The accelerometers were mounted radially and vertically with respect to the pinion. The magnetic tachometer was installed above the right pinion and measures speed from the pinion pulses as the tooth passes under the sensor. The gear ratio was used to process the data at the correct speed for the gear.

In the test rig, a second parallel data acquisition system referred to as the Mechanical Diagnostic System Software (MDSS), was also used to measure accelerometer, tachometer and torque sensor data. The MDSS accelerometers were installed in the same orientation in close proximity to the MSPU accelerometers on the gearbox housing. These accelerometers had a 0.7 to 20 KHz frequency range with a resonant frequency of ≥70 KHz. Speed was measured with optical tachometers mounted on both the left pinion shaft and left gear shaft, producing separate once-per-revolution tachometer pulses for the pinions and gears. The location of the MSPU and MDSS accelerometers in the test rig are shown in Figure 3.

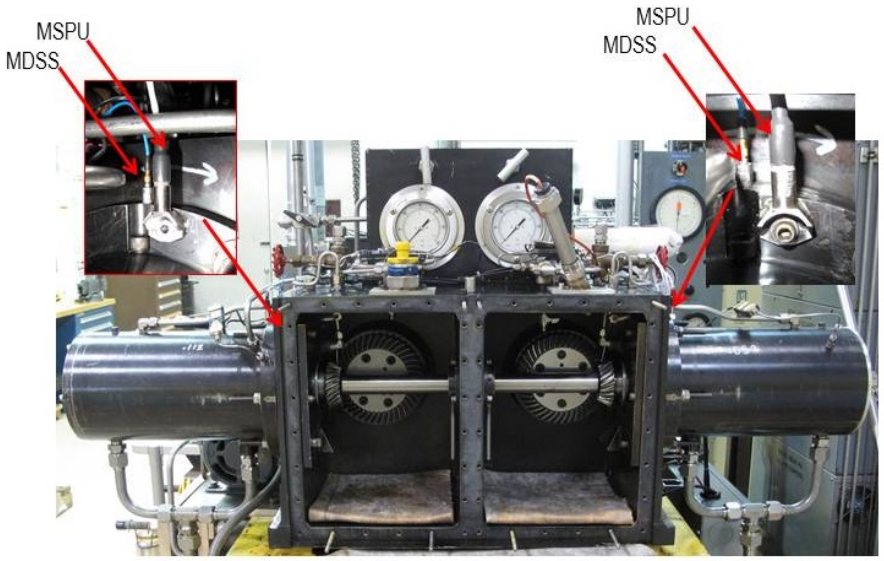

Figure 3.—Location of MDSS and MSPU Accelerometer

TABLE 1.—GEAR FAILURE MODES (REF. 12)

Class	Mode	Degree
Contact Fatigue	Subcase Fatigue	
	Micropitting	
	Macropitting	Initial
		Progressive
		Flaking
		Spalling
Scuffing	Scuffing	Mild
		Moderate
		Severe

5.0 Failure Modes

The failure modes investigated were defined by class (contact fatigue), general mode (macro pitting) and degree (progressive) per American Gear Manufacturers Association (AGMA) standards for gear wear terminology (Ref. 12). A pitting fatigue failure mode is when material is expelled from the gear tooth contact. Scuffing was also observed on some gear teeth. Scuffing is not a fatigue failure mode, but rather the transferring of metal due to welding and tearing caused by insufficient elastohydrodynamic oil film thickness. Table 1 illustrates the types of damage observed. Definitions for tooth surface macro pitting modes (Ref. 12) are summarized as follows:

- Initial—Pits less than 1 mm in diameter.
- Progressive—Pits in different shapes/sizes greater than 1 mm in diameter.
- Flake—Pits that are shallow thin flakes.
- Spalling—Pits that cover tooth contact surfaces that exceed progressive pitting.

It's interesting to note that like these failures, a 30 year survey of helicopter investigations found 55 percent of component failure was due to fatigue type failure modes described as pitting and spalling while fretting or scuffing was found as the third highest factor (Ref. 13). The majority of these failures,

over 34 percent, were found in the engine and transmission. Fatigue failures found on gear and pinion teeth were found to initiate due to alignment, manufacturing defects and lubrication issues (Ref. 13).

6.0 System Factors

Gear vibration based CIs are based on vibration signatures created at the gear mesh as the pinion and gear teeth come into contact. The vibration signatures change when the teeth are damaged. These vibration signatures are measured by an accelerometer installed on the gear housing. As discussed in previous reports, many factors can affect a vibration based condition indicator's ability to respond to tooth damage. System factors between the helicopter and test rig that can affect CI response are highlighted in Figure 4. Comparable system factors are identified in green boxes. Factors with some differences are identified in yellow boxes. Factors with significant differences are identified in red boxes.

The fault type (gear or pinion), mode, class, degree, magnitude and how it initiates and progresses affect the measured vibration response. The fault also interacts with the rig and gear design based on the path to the accelerometer. The failure modes observed in the test rig and on the helicopter were comparable for both test rig and the helicopter gearbox. However, damage progression to the gear teeth was documented with photographs during inspections throughout a test, while only the damage of the gear teeth at replacement was documented.

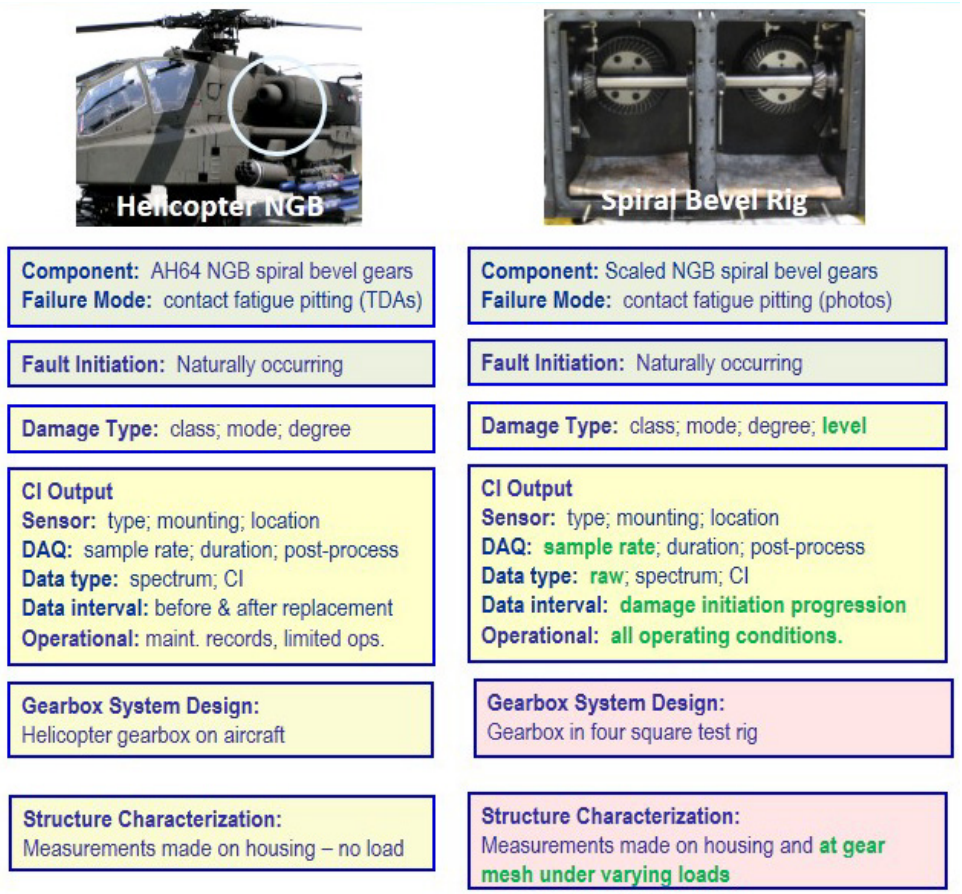

Figure 4.—Factors that affect Helicopter and Test Rig CI response

The response of the accelerometer to a specific fault can depend on the sensor specifications, mounting, signal processing of the raw signal and the method of CI calculation. These factors were the same since the same HUMS was used for both systems. They differed in that additional data was measured and stored for the test rig that included raw time domain vibration data and operational data, such as torques, temperatures and oil pressures. Due to the limited on-board storage capability, raw data and operational data was not stored on the helicopter HUMS.

Several other parameters, such as sensor location, operational conditions and drive train design can affect CI response. These parameters were significantly different between the helicopter and the test rig. The location of the accelerometer on the housing with respect to the gear and pinion was different due to differences in gearbox designs. Operationally, rig tests were run at speeds significantly lower than the helicopter. The four square rig design, in which the left pinion drives the gear and the right gear drives the pinion, and the coupling between the right and left gearboxes were significantly different than the helicopter. All of these parameters have the potential to cause differences between helicopter and test rig vibration response. In an attempt to better understand these differences, a modal analysis was performed on the helicopter gearbox and test rig to characterize the two systems and shed some light on what differences may be expected. These measurements will be discussed in Section 9.0, Structural Characterization.

7.0 Condition Indicators

Since current helicopter gears generate vibration signatures synchronous with gear rotational speed, all gear condition indicators are calculated from time synchronous averaged data (TSA). TSA data contains vibration signals averaged over several revolutions of shaft speed, measured by the tachometer, to improve signal to noise ratio (Ref. 14). Vibration signals synchronous with the shaft speed intensify relative to non-periodic signals. Using the tachometer, the vibration signal collected from the accelerometer is interpolated into a fixed number of points per shaft revolution. Both the MSPU and MDSS data acquisition systems sampled the vibration data at speeds that provide sufficient vibration data for calculating asynchronous and time synchronous averaged data based on the component rotational speed.

For this investigation, four gear condition indicators were compared and include Figure of Merit 4 (FM4), Root Mean Square (RMS) referred to as DA1 in the MSPU system, Sideband Index averaging ± 1 sideband (SI1), and Sideband Index averaging ± 3 sidebands (SI3) referred to as SI in the MSPU system (Refs. 7, 14, 15, and 16). Note that SI1 is not currently calculated in the on-board MSPU, rather from post-processed TSA data. Figure 5 shows block diagrams of the steps required to calculate the condition indicators. More details of these condition indicator calculations for the rig and helicopters can be found in References 1 and 2.

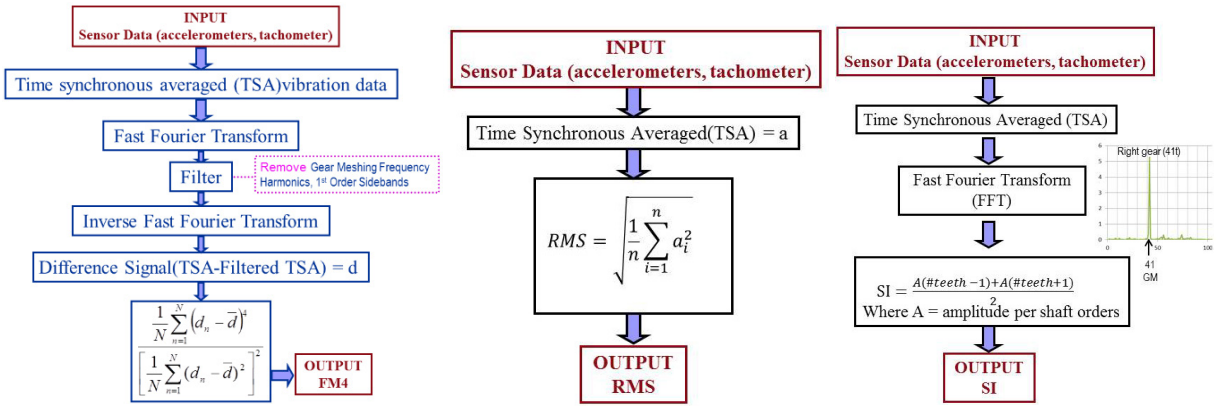

Figure 5.—FM4, RMS and SI Condition Indicators

8.0 Summary of Helicopter and Test Rig CI Response

A subset of the helicopter and rig data will be compared for this analysis. Prior to comparison of the CI response from both systems, data sets from both are summarized. For rig data, tables were generated for each test, identifying the failure modes observed on the gear teeth at inspection intervals and color coded based on damage mode. For the helicopter, tables were generated for the helicopters identifying the damage observed during tear down of the gearbox when it was replaced.

Table 2 contains a summary of the test rig data for a total of eight tests. Damage modes and levels were determined by visual inspection and examination of inspection photographs of all the damaged teeth, a total of 120 pictures for each inspection, per inspection interval. The first two columns of each table indicate the number and inspection interval in minutes. The next four columns indicate the four components investigated: gear left (GL), gear right (GR), pinion left (PL), pinion right (PR) and the number of teeth with damage observed at each inspection interval. The color coded damage scale is shown in a table in the bottom right. The test number is shown in the first row of each test inspection table. Comparable failure modes are color coded for further analysis. For example, for tests L4545R5050 and L1515R5050 the left pinion was damaged at the end of the test.

TABLE 2.—INSPECTION TABLES AND DAMAGE SCALES FOR RIG TESTS

Inspection	L4545R5050				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 76				
3	76 to 324				
4	324 to 1370				
5	1370 to 2120			1	
6	2120 to 2403			2	
7	2403 to 2833			2	

Inspection	L3030R5050				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 70				
3	70 to 1784			micro	
4	1784 to 3270			micro	
5	3270 to 4633	1		micro	
6	4633 to 5359	1		micro	
7	5359 to 5962	2		1	
8	5962 to 6037	2		1	

Inspection	L1515R5050				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 63				
3	63 to 705			1	
4	705 to 1022			2	
5	1022 to 1291			2	

Inspection	L2020R5050				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 70				
3	70 to 217	s-all		s-all, pit	

Inspection	L4040R5050				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 63				
3	63 to 370	s-all		s-all	

Inspection	L3535R5050				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 178				
3	178 to 636				
4	636 to 6276				
5	6276 to 6818	2			
6	6818 to 7617	2			
7	7617 to 9358	2		1	
8	9358 to 9578	2		1	

Inspection	L1818R1616				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 277				
3	277 to 546				
4	546 to 3603				edge
5	3603 to 5249			1	edge
6	5249 to 7065			1	edge
7	7065 to 9062	4		3	edge

Inspection	L2121R1919				
1	Inspection (min)	GL	GR	PL	PR
2	1 to 127				
3	127 to 307	s-all		s-all	edge
4	307 to 1122	s-all		5	edge
5	1122 to 1393	s-all		6	edge
6	1393 to 1568	s-all		8	edge
7	1568 - 1905	4		10	edge

Damage Scale
No damage
micro pitting or edge wear
scuffing
1 tooth macropitting
> =2 teeth macropitting

Table 3 is a summary of the helicopters analyzed with damaged spiral bevel gear sets. The first column identifies the tail number. The second column identifies if the damaged gear set was found in the left or right nose gearbox. The next four columns indicate the four components investigated: pinion right (PR), gear right (GR), pinion left (PL), gear left (GL). An x in the column indicates damaged was observed on this component prior to replacement. A summary of the damage described in the tear down analysis report is listed in the last column.

Table 4 and Table 5 summarize the performance of the CIs to specific component faults for the helicopter and the test rig. These tables were generated by reviewing the CI data from helicopters and rig tests detailed in References 1 and 2. The CI response of the helicopter data is compared per tail and the CI response of the test rig data is compared per test. A good helicopter CI response is identified when the CI crossed over a discrete threshold prior to replacement and decreased after replacement. A good CI response for the test rig is when the CI crossed over a threshold as damage was observed on the gear teeth during inspections. The thresholds to assess performance are shown in the top row of both tables. Tails with comparable failure modes and faulted components are color coded the same. The color codes are listed in Table 6. Note that the scuffing failure mode, after run-in, could only be identified for rig tests. The minimal documentation available for the helicopters made it impossible to isolate this failure mode. A threshold table generated in Table 7 illustrates the differences in thresholds between the helicopter and test rig for DA1, SI1 and SI3. DA1, for example, was a 23X higher on the helicopter than the test rig. One cause of these differences will be discussed in Section 9.0, Structural Characterization.

In Table 4 and Table 5, the first column identifies the tail number or test number. The next four columns indicate the four components investigated: right pinion (PR), right gear (GR), left pinion (PL), left gear (GR). An x in the column indicates damaged was observed on this component. The cells highlighted in green indicate the CI of the damaged component responded to damage. The cells highlighted in blue indicate the meshing component responded. The cells highlighted in light grey were unresponsive. For the helicopter, the "*" indicates the CI responded during the 120 knots true airspeed (120KTA) regime. If no "*" is shown, it indicates the CI measured during the 101 percent rotor speed flat pitch ground (FPG101) regime.

TABLE 3.—SUMMARY OF HELICOPTERS WITH SPIRAL BEVEL GEAR DAMAGE

#	Component	PR	GR	PL	GL	Damage
1	Left			x	x	Gear and pinion teeth uneven wear and *scuffing*; output spline wear
2	Left			x	x	Gear and pinion severely damaged
4	Left			x		Gear wear and pinion pitting; output bearing IR spalling
5	Left			x		Pinion pitting; spline damage
6	Right	x				Moderate wear and *scuffing* on pinion
7	Right	x	x			Gear and pinion wear and pitting; gear scuffing
8	Right	x	x			Gear and pinion teeth *scuffing*; input bearing roller spalling
10	Right	x				Pinion moderate scuffing; output duplex *bearing* ball spalling
13	Right	x	x			Gear tooth chipped on toe side of gear, gear and pinion *scuffing*, pinion spalling
14	Left			x	x	Gear and pinion spalling
15	Left			x	x	Gear and pinion *scuffing* and spalling
16	Left			x	x	Gear and pinion severe *scuffing* on edge of teeth
18	Left			x		Pinion severe spalling
20	Right	x	x			Pinion and gear severe pitting

TABLE 4.—SUMMARY OF HELICOPTERS WITH SPIRAL BEVEL GEAR DAMAGE

Helicopter								
#	PR	GR	PL	GL	DA1 ≥ 80	FM4 ≥ 4.5	SI1 ≥ 6	SI or SI3 ≥ 6
1			x	x		GL		
2			x	x				
4			x		PL,PL*,GL,GL*		PL,PL*,GL,GL*	PL
5			x				PL,PL*	PL*
6	x						PR*	PR*
7	x	x				GR		
8	x	x					PR,PR*	
10	x						GR*	
13	x	x			PR,PR*	GR*	PR,PR*	PR,PR*
14			x	x	PL,PL*,GL,GL*		PL,PL*,GL,GL*	PL,PL*,GL,GL*
15			x	x	PL		PL	PL
16			x	x	PL,PL*,GL,GL*		GL,PL	
18			x			PL		PL
20	x	x			GR,PR		GR,PR	GR

TABLE 5.—SUMMARY OF RIG TESTS WITH SPIRAL BEVEL GEAR DAMAGE

Test Rig								
Test	PR	GR	PL	GL	RMS ≥ 3.5	FM4 ≥ 4.5	SI1 ≥ 0.5	SI or SI3 ≥ 0.35
L4545R5050			x		PL	PL	PL	PL
L3030R5050			x	x		GL	PL	PL
L1515R5050			x			PL		
L2020R5050			x	x	GL, PL		PL	
L4040R5050			x	x	GL, PL			
L3535R5050			x	x	GL, PL	GL	GL, PL	PL
L1818R1616	x		x	x			GL,PL	PL
L2121R1919	x		x	x	GL, PL			

TABLE 6.—COMPONENT AND
DAMAGE MODE COLOR CODES

Damage Modes	
	Gear & pinion pitting, spalling or scuffing
	Pinion only pitting, spalling or scuffing
	Scuffing after run-in

TABLE 7.—COMPARISON OF TEST RIG AND HELICOPTER THRESHOLDS

Thresholds	DA1 ≥ T	FM4 ≥ T	SI1 ≥ T	SI3 ≥ T
Helicopter (T)	80	4.5	6	6
Test Rig (T)	3.5	4.5	0.5	0.35
Factor	23	1	12	17

After reviewing Table 4 and Table 5, some interesting inferences can be made in regards to the two systems for each damage case. For the pinion and gear damage case, DA1/RMS detection rate was better in the helicopter than test rig due to the test rig sensitivity to operational conditions discussed in Reference 1. FM4 performed better in the test rig than helicopter and was detected by the gear CI, not the pinion, for both the helicopter and the test rig. SI1 performed the best at detecting gear and pinion damage for both the test rig and the helicopter.

For the pinion only damage case, SI1 and SI3 performed better in the helicopter than the test rig while FM4 performed better in the test rig than the helicopter. For both damage cases, the thresholds for DA1/RMS, SI1 and SI3 were significantly higher in the helicopter than the test rig. Since FM4 is normalized, the thresholds were the same for both the test rig and helicopter. For the test rig scuffing damage case, RMS responded to this failure mode.

References 1 and 2 contain plots of the CI response for the helicopter and the test rig for FM4, SI1, SI3 and DA1. Figure 6 to Figure 11 are representative plots comparing test rig to helicopter CI data for two rig tests and two helicopters at FPG101 regimes. Figure 6 and Figure 7 compare FM4 and SI or SI3 from rig test L4545R5050 to helicopter tail #18. Figure 8 to Figure 11 compare FM4, SI1, SI3 and DA1 or RMS from rig test L3535R5050 to tail #14. For the test rig, Table 2 identifies the color coded damage scale at inspections and is noted on the X-axis by color coded triangles. For the helicopter, the green square indicates when the gearbox was replaced. Refer to Table 4 and Table 5 for indications of CIs that responded to damage cases for the test rig and helicopter. Both MDSS and MSPU component CI values are plotted for the test rig.

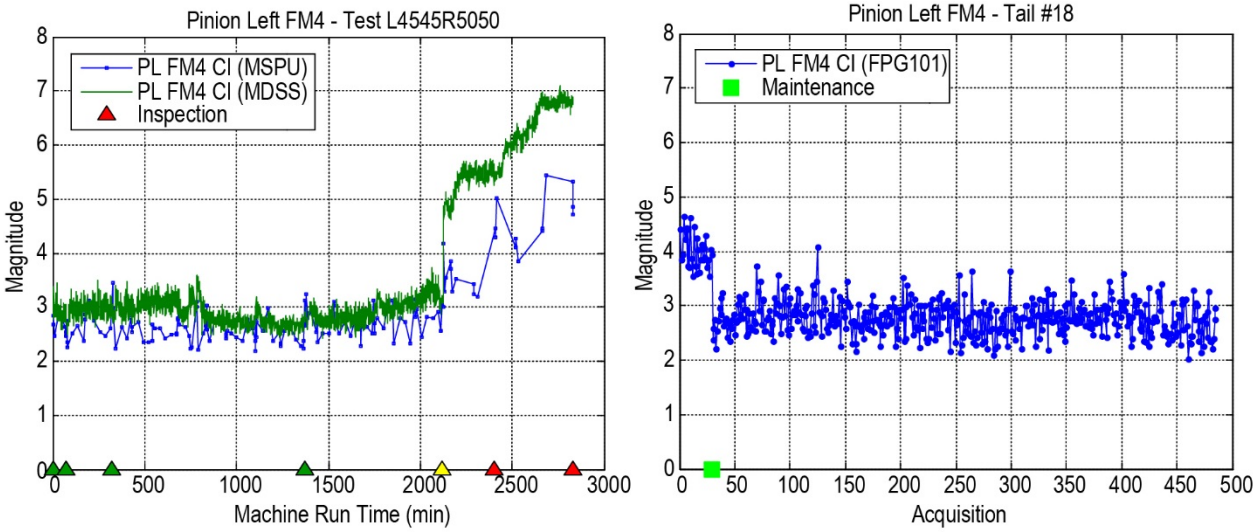

Figure 6.—Comparison of Helicopter # 18 and Test Rig L4545R5050 FM4 Response

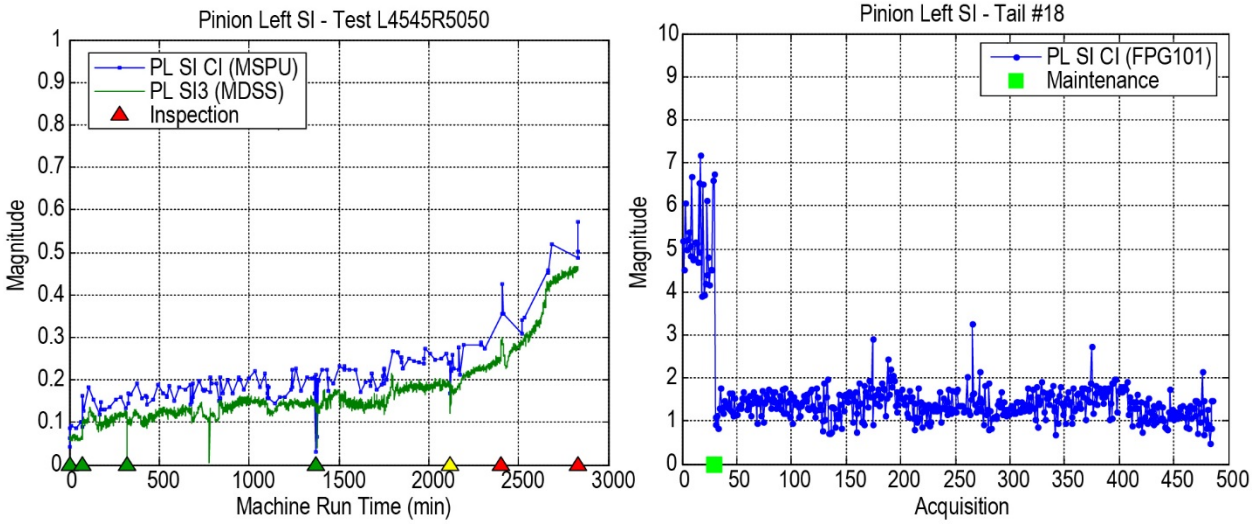

Figure 7.—Comparison of Helicopter # 18 and Test Rig L4545R5050 SI Response

Figure 6 and Figure 7 compare tail #18 and test L4545R5050 for the pinion only the damage case. Both figures show FM4 and SI/SI3 responded for both the test rig and helicopter. It also shows the threshold for SI/SI3 was significantly higher in the helicopter than the test rig.

Figure 8 to Figure 11 compare tail #14 and test L3535R5050 for the pinion and gear damage case. Figure 8 shows FM4 responded on the test rig, but not on the helicopter. Figure 9 to Figure 11 show SI1, SI3 and DA1/RMS responded for both the test rig and helicopter. It also shows the thresholds for SI1, SI3 and DA1/RMS were significantly higher in the helicopter than the test rig. The sensitivity of DA1/RMS in the test rig to torque is also shown in Figure 11.

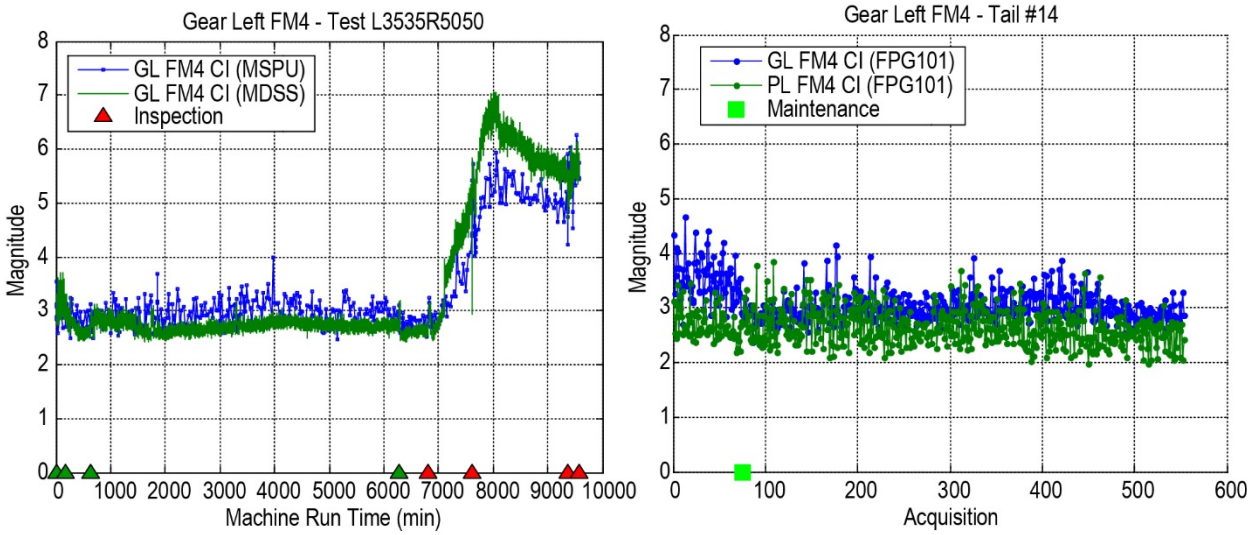

Figure 8.—Comparison of Helicopter # 14 and Test Rig L3535R5050 FM4 Response

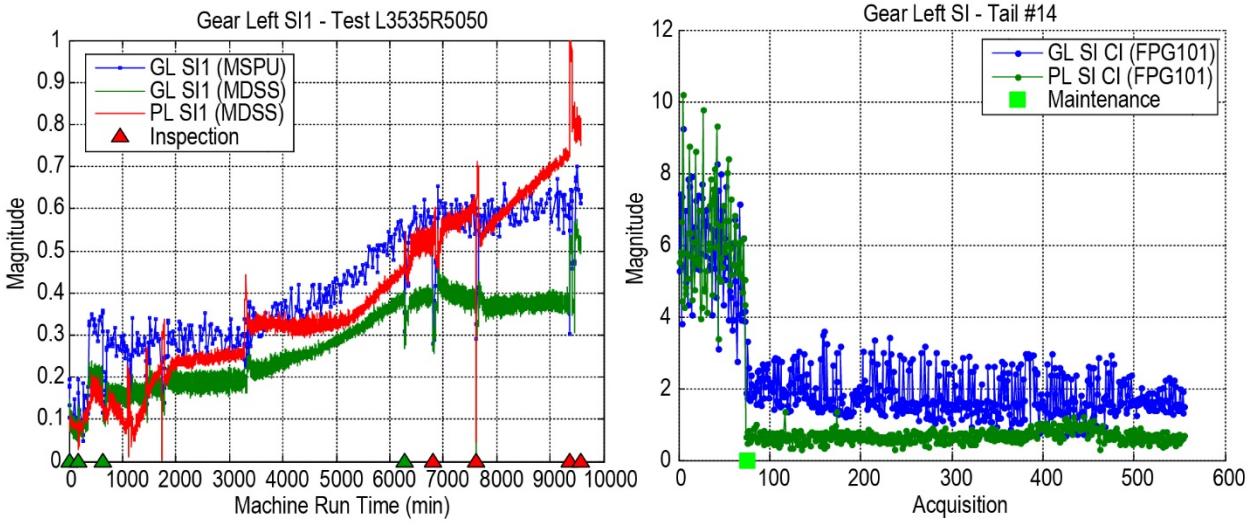

Figure 9.—Comparison of Helicopter # 14 and Test Rig L3535R5050 SI1 Response

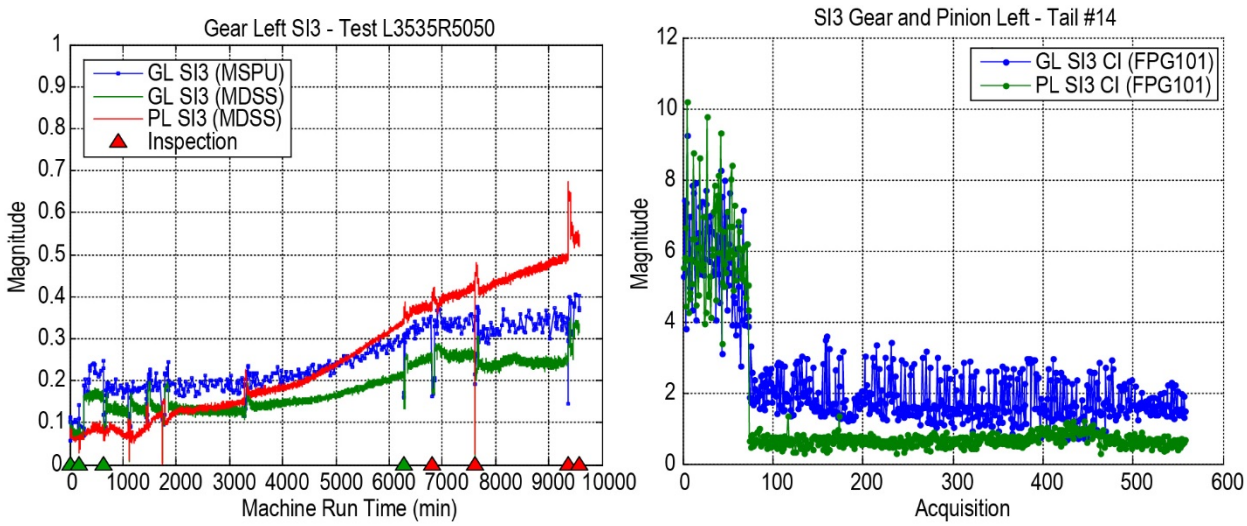

Figure 10.—Comparison of Helicopter # 14 and Test Rig L3535R5050 SI (SI3) Response

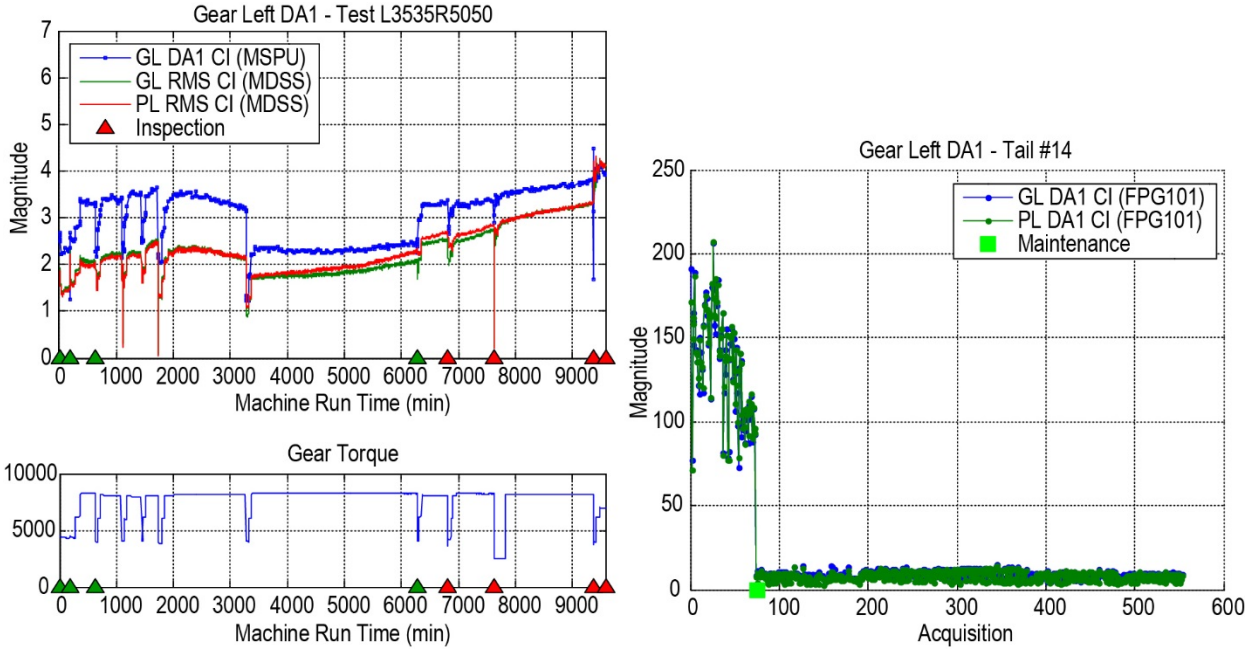

Figure 11.—Comparison of Helicopter # 14 and Test Rig L3535R5050 SI Response

9.0 Structural Characterization

HUMS use vibration signatures generated from damaged components to identify transmission faults. Their performance can be affected by the transfer path of this signature through the structure. The structure can filter the response between the vibration source and the housing mounted accelerometer by its design and sensor location. Structural differences between a component test rig, a full-scale transmission test stand or an aircraft may be one cause of differences in CI performance when measured on different systems. For this reason, measurements were made to characterize helicopter and test rig static structural dynamics.

The purpose of these measurements were to better understand differences between vibration data measured on test rig and helicopter and how these differences may direct testing for CI development. Measurements were made on a helicopter gearbox, on a gearbox removed from the helicopter and installed in a NASA test fixture and on the Spiral Bevel Gear Fatigue Test Rig gearbox. As shown in photographs of the three geared systems, in Figure 12, the gearbox housings are very different, indicating the vibration transfer paths between the fixture and the test rig would also be different.

A modal analysis was used to characterize the three systems. A modal analysis is performed by exciting a structure with a known impact, and measuring the vibration response with an accelerometer located on the structure. Frequency response function (FRF), defined as a measured response normalized by measured input, are used to analyze structural dynamics under static conditions. FRF measurements taken on the three systems were excited with an instrumented impact hammer or a commercial reaction-mass shaker in the form of a piezoelectric actuator. The response was measured with the accelerometers installed at the same locations they were installed for CI measurements for the helicopter and NASA NGB test fixture.

Measurements on the helicopter mounted nose gearbox were made by the U.S. Army. These measurements were made on the helicopter to help select ideal bearing defect detection frequency bands, not to better understand gear CI performance (Ref. 17). These measurements were not taken under varying load conditions and were only made on the external housing of the aircraft.

The U.S. Army provided NASA Glenn with a helicopter nose gearbox for additional evaluation of the frequency response. A test fixture that provides a static torque to the input side of the gearbox and output gear locked to ground was developed. In addition to repeating the external housing measurements, measurements were also taken impacting at the gear mesh to simulate gear dynamics. This was done because gear condition indicators use signatures related to changes in dynamics due to the meshing of damaged teeth.

Frequency Response functions were compared using the Kolmogorov–Smirnov (KS) statistical test. The KS test is a nonparametric statistical test for comparing probability distributions that makes no assumption about the distribution of the data (Ref. 18). The KS statistic, D, quantifies a maximum vertical distance between the empirical cumulative distribution functions (CDF) of two samples. KS Similarity factors were also calculated by subtracting D from 1.

First, the helicopter and fixture mounted FRFs were compared. External measurements taken on the gearbox in the fixture were shown to be consistent with those taken on the helicopter under the same conditions. Measurements were then made in the fixture with the shaker on the gear at several loads (Ref. 19). These measurements were made on the gear because the transfer path is very different from that measured on the gearbox housing.

Figure 12.—Three spiral bevel geared systems

Measurements were then taken on the Spiral Bevel Gear Fatigue Test Rig under conditions comparable to the nose gearbox fixture measurements. Details of these measurements can be found in Reference 20. The most comparable measurement taken on both systems was the installation of an actuator on the ring gear while measuring vibration response at the housing. Figure 13 contains photographs of actuators installed in the test rig gear and the nose gearbox fixture. FRF and corresponding CDF plots are shown in Figure 14. The top figure FRF amplitude is plotted on a log scale. The KS Similarity factor for 0 to 14 kHz was equal to 0.10, which indicates very different vibration transfer paths between the two structures. This is no surprise based on the significantly different gearbox designs of the two systems. One interesting observation was that the FRF amplitude measured on the helicopter gearbox was significantly higher than the test rig.

In the helicopter, the left and right nose gearboxes are separated by many other shafts and gears within the aircraft transmission. In the test rig, the left and right gear sets are only separated by one shaft and reside in the same gearbox housing. Measurements were made to better understand how the four square design of the test rig affects CI response. The test rig was designed to test two sets of spiral bevel gears simultaneously. Due to the cross shaft between the pinions on the left and right side, both gear sets are directly coupled. The cross shaft connecting the right and left pinions can make it more challenging to isolate tooth damage detection to one side of the gearbox and separate CIs originating at either gear set.

To investigate how damage on one side can potentially affect the other, impacts were made at the left pinion when 100 percent preload was applied, while the FRF was measured on the right and left side. FRF and corresponding CDF plots are shown in Figure 15. The top figure FRF amplitude is plotted on a log scale. A KS similarity value of 0.65 was calculated for the 0 to 6400 Hz frequency range. This indicates some coupling between the right and left gearboxes through the shaft. The unique design of the test rig causes difficulties in isolating pinion tooth damage on one side of the gearbox.

Several important things were learned from the modal analysis. The first was that the transfer path dynamics are about an order of magnitude higher on the NASA NGB fixture gearbox than on the test rig. This means that for a given forcing function, such as changes in the vibration response at the gear mesh, higher response amplitudes would be measured by the accelerometer on the helicopter gearbox. Based on this measurement alone, it would be harder to detect defects on the test rig than on the helicopter. But wait, there's more. Since the gears turn much faster on the helicopter than on the rig, the defect forcing function would be higher on the helicopter as well, further indicating a CI would be more responsive to defects in the helicopter gearbox.

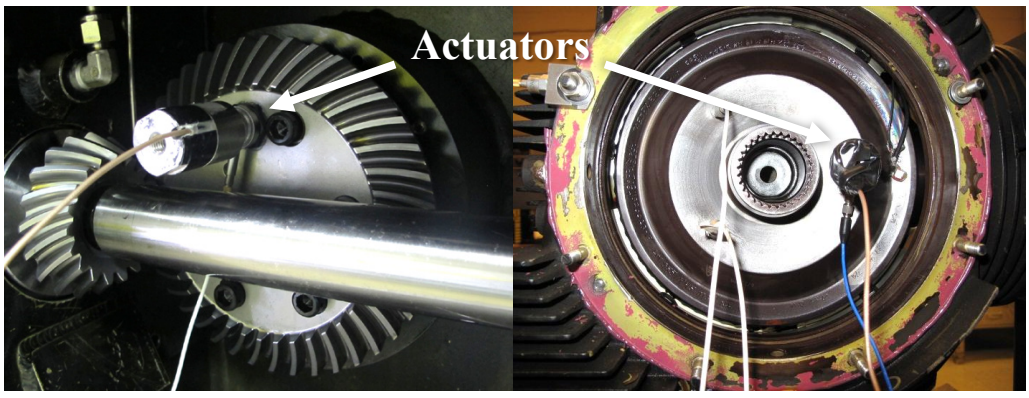

Figure 13.—Actuators installed on the test rig and nose gearbox fixture.

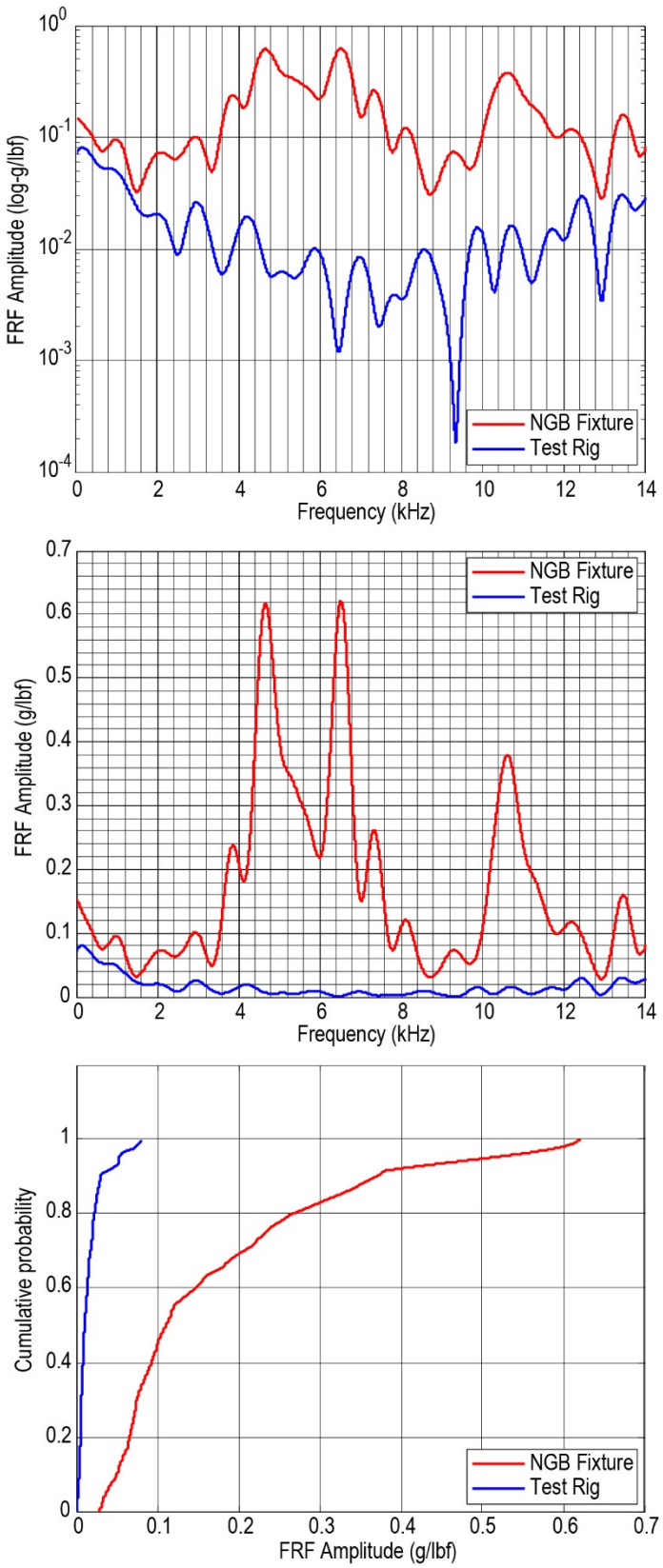

Figure 14.—FRF and CDF of NGB fixture and test rig

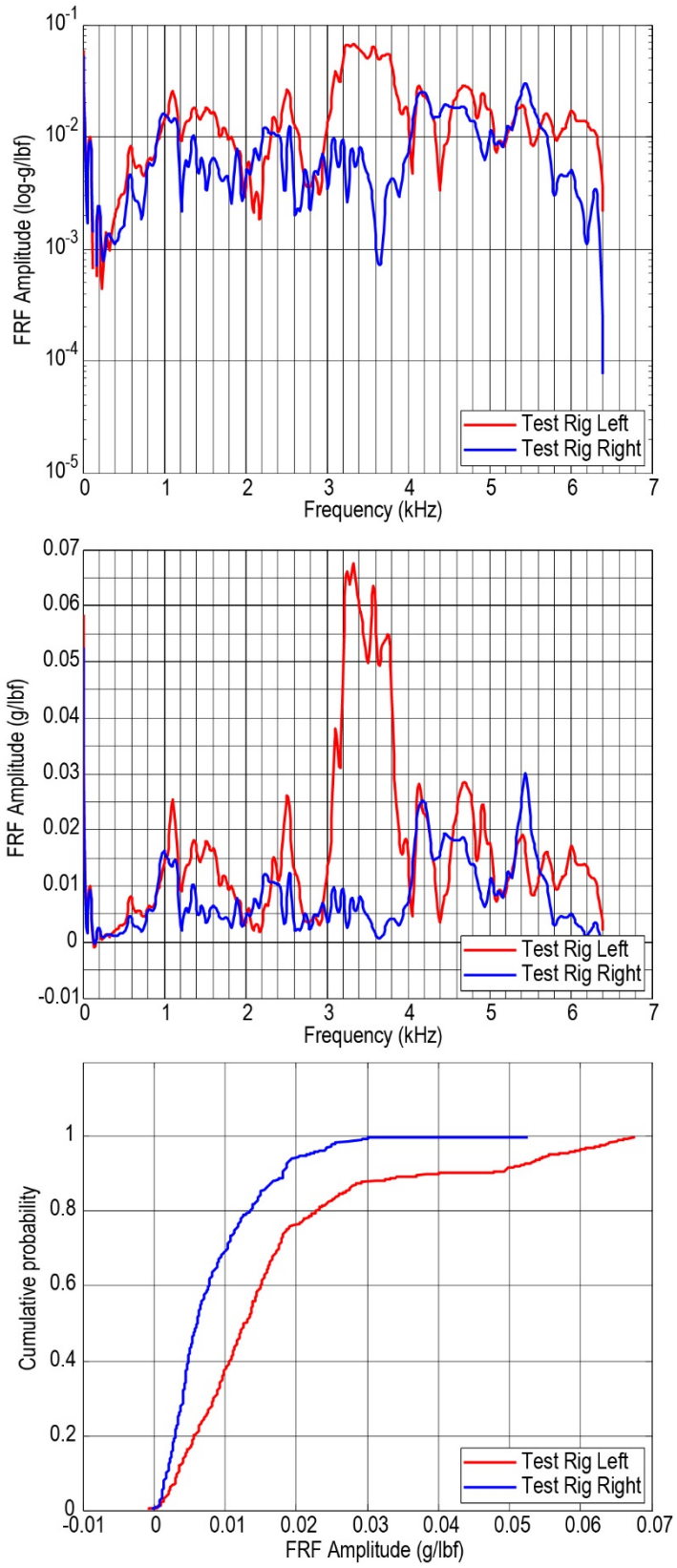

Figure 15.—FRF and CDF of left and right test rig gearbox

The test rig design also had an effect on vibration response. Response from the left and right side of the test rig are difficult if not impossible to separate—this is not the case on the helicopter. This has implications for running two gear sets with defects. For this reason, super-finished gears were installed on the right side of the test rig in an attempt to force a failure on the left side. This was not always the case for all the tests. And, even if only one gear set was damaged, if gear-tooth loading/unloading does not happen at exactly the same time for all speeds and loads, it can also affect the vibration signature.

Gear design can also cause more issues when measuring vibration response in the test rig. Slight manufacturing differences within each gear set can cause variances in the gear meshing frequencies on the right and left side. Differences in the relative phase of the right and left gear mesh frequencies may be higher or lower than only one rotating gear set. These condition have the potential to smear some of the spectral peaks generated by the defect. Results of these measurements indicate gear tooth damage will be more challenging to detect in the test rig than the helicopter.

Other environmental conditions such as operating conditions can also affect measured vibration response. Speed sweeps were performed on the test rig to identify resonances and specific operating speeds. Tests were performed at speeds that did not coincide with a resonance to eliminate this effect on vibration response.

10.0 Test Rig CI Validation

Understanding the complete monitored system is important before attempting to interpret health monitoring data. Figure 16 identifies many of the system factors that can affect the measured vibration response when pinion and gear teeth come into contact. A reliable gear condition indicator must be tuned to the specific fault vibration signature with minimal effect from other factors. Dissimilarities in systems design and environment under dynamic operating conditions caused differences in CI response between the helicopter and test rig. The factors, in the box, highlighted in green are those that were maintained between the helicopter and the test rig. The yellow highlighted factors indicate the majority of the factors remained the same except those listed in red font. The factors highlighted in red identify parameters that were significantly different.

The same HUMS, referred to as MSPU, was used to measure CI data on both the helicopter and the test rig. The data acquisition system, CI calculation methods and accelerometers were the same. However, the location and mounting were different due to the differences in gearbox design. An accelerometer must be mounted at a location sensitive to the component frequencies under investigation since its mounting and location can have large effects on the measured vibration response. Location and mounting can be optimized to obtain the best response, although installation is often limited to space availability on the helicopter.

In regards to the test rig and gear set design, the gears had a diametral pitch different than the helicopter gear set. Profile and spacing errors varied slightly from tooth to tooth due to design tolerances. These manufacturing differences can also occur in the helicopter. However, due to the four square rig design, these minor differences between gear sets had a significant effect in the test rig.

Differences in operating speeds can produce different vibration source amplitudes due to imbalance and other component tolerances within the system. Plus, if the CI is sensitive to environmental conditions, parameters such as torque and speed must be measured while maintaining steady flight regimes. For these two systems, the test rig proved more challenging to detect defects than the helicopter due to its sensitivity to operational and environmental conditions. Keep in mind that for some fielded components, additional noise sources found on the helicopter could overwhelm defect frequencies.

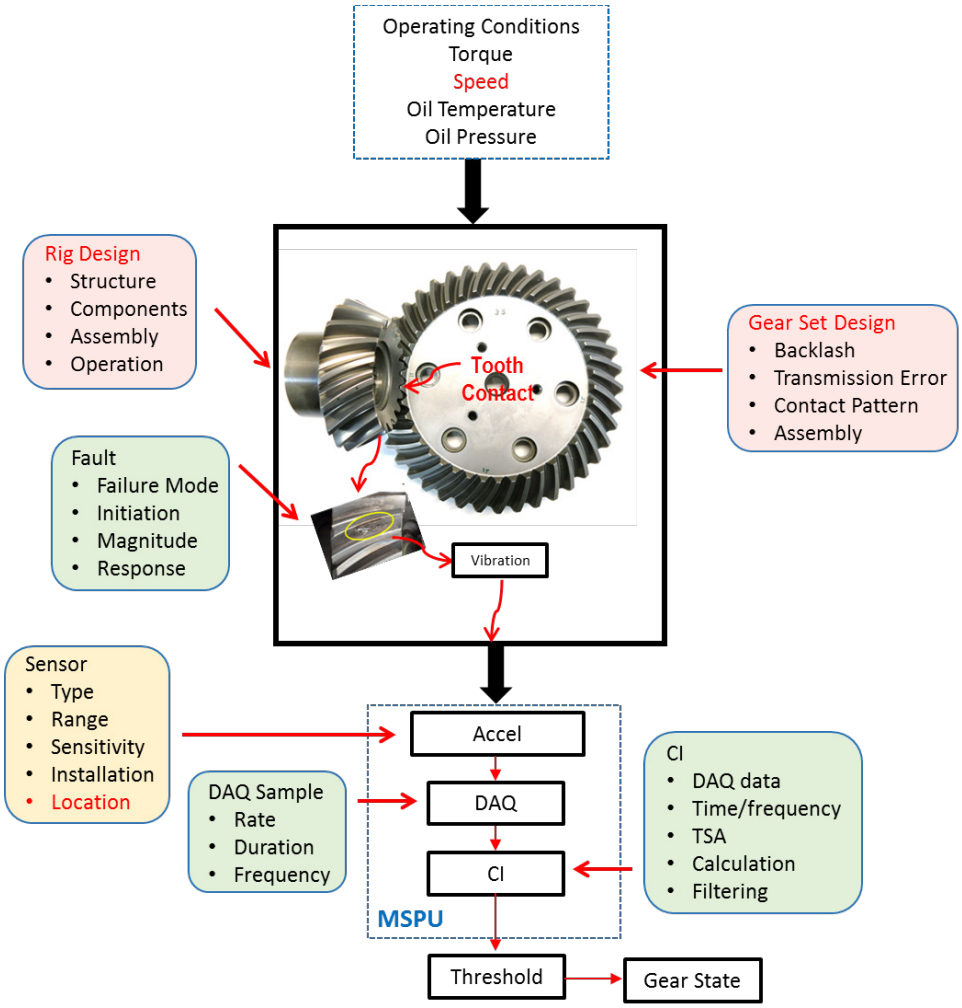

Figure 16.—System Factors

A CI with minimal effect from other factors such as gear and system design, system maintenance, sensing system and data acquisitions methods while only responding to gear failure modes are desired CI characteristics. Due to differences in both systems and their operational environments, test rig CI response was not always representative of helicopter CI response. For these situations, CI performance limitations must be defined to understand the risks in using a test rig validated CI on a helicopter.

Figure 17 provides a process flowchart for the steps required for test rig condition indicator validation. Reviewing Figure 17, once the platform or aircraft, gearbox, mechanical component, failure mode, physics of the failure mode (dynamics that change due to fault) and HUMS are identified, the next step is to identify a test rig. If the platform gear set is not used, the design method and assumptions must be documented. If a different gearbox is used, the effect of the structure on vibration response must be quantified through a modal analysis.

Even if the same gearbox and gear sets are used in the test rig, mounting of the gearbox and supporting systems in a test facility that vary from the helicopter may also affect vibration response. A means to evaluate the effect of the supporting hardware design on vibration response must be understood. One method would be to perform a speed sweep at operating conditions and verify no resonances around the frequencies of interest. If torques and speeds on the helicopter cannot be maintained in the test rig, their effect on vibration response must be quantified.

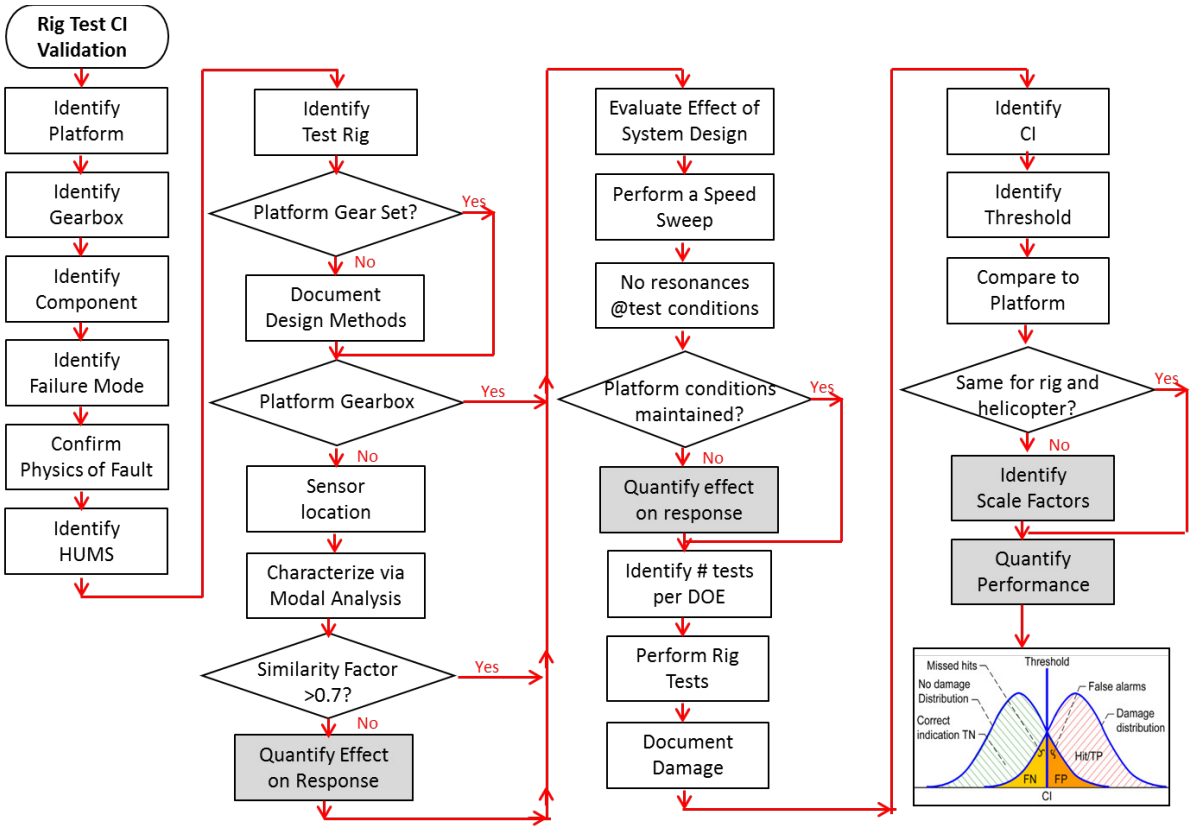

Figure 17.—Process Flowchart for Rig Test CI Validation

The number of gear sets to test must be quantified by a design of experiments (DOE) approach. This requires a hypothesis defined from the questions the experimental investigation are trying to answer. A test statistic must also be defined about the data set. This method determines the number of samples required to evaluate the hypothesis. Details on sample size determination can be found in Reference 21.

Once rig testing begins, a process to document damage and progression must be implemented. This documentation is required to confirm the correct failure mode is generated in the test rig and to eliminate a subjective interpretation of damage. The failure mode must also be correlated to the CI response through the time of the inspection and the time stamp of the HUMS data. A threshold is identified for damage indication. The threshold can be defined by the CI value crossing over a discrete limit, or a rate change, or a change in statistical distributions, or a hypothesis test or by a rule based expert system or through an automated reasoner.

The baseline data from the test rig must be compared to the baseline from the helicopter. If damaged data is available from the helicopter, this can also be compared. If the CI values and threshold differ significantly, a way to scale them must be identified. This will require using the methods for quantifying the effect of the system on vibration response. These steps are highlighted in grey.

The final step is to quantify CI performance metrics using damage detection and false alarm rates. The final block provides a visual graph of two normal distributions used to represent a no damage response and a damage response of a CI. The threshold line separates the graph into correct indication/TN-true negative (no damage—no indication), FN-false negative (damage present—no indication), false alarms/FP-false positive (no damage—indicated) and hits/TP-true positive (damage—indicated). Table 8 also illustrates the system states from CI response and parameters for calculating damage detection rates, false alarm rates and the accuracy of the CI using a confusion matrix (Ref. 22).

Referring to Table 8, the damage detection or true positive rate (TPR) is equal to a/(a+c). The false alarm or false positive rate (FPR) is equal to b/(b+d). The accuracy of the CI at correctly indicating damage and healthy states is equal to (a+d)/(a+b+c+d). Aligning HUMS data to make these calculations can be challenging if the true state of the gear teeth is unknown. These calculations were made for the test rig MDSS data using the thresholds listed in Table 7. The test rig MDSS data was used instead of the MSPU data because significantly more data was available from this system due to the higher acquisition rate. An example of the steps required to make these calculations will be discussed for one rig test, L4545R5050 and is outlined in Table 9.

Table 2 provides the inspection intervals and damage scales for test L4545R5050. The calculations are limited to the damage mode of macro pitting on two or more teeth. The no-damage state data used is in the time frame listed in green from 1 to 1370. The damage state data used is in the time frame listed in red from 2120 to 2833. The data within the yellow region will not be used since this is not the damage mode of interest and this region was also the transition region, when the damage state went from green to red. In Table 9, the first column identifies the number of cases, or data points, available within each inspection interval for each gear state. The second column identifies the observed state at that inspection. For this test, only the left pinion teeth had damage. The next four columns identify the number of times the CI values exceeded the thresholds within the inspection interval.

TABLE 8.—STATE OF SYSTEM – HEALTH OF COMPONENT

CI Response	Damage	No Damage	Total
Damage Indicated	True Positive (TP) Hits, #detected faults = a	False Positive (FP) False Alarms, #false alarms = b	Number of alarms a+b
Damage not Indicated	False Negative (FN) Missed Hits, #missed hits = c	True Negative (TN) Correct Indication, #correct rejections = d	Number of no alarms c+d
Total	Number of faults a+c	Number of healthy b+d	Number of cases a+b+c+d

TABLE 9.—CALCULATING TPR, FPR AND ACCURACY FOR RIG TEST L4545R5050

L4545R5050	Thresholds	3.5	4.5	0.5	0.35
Cases	Pinion Left CI	PL RMS CI	PL FM4 CI	PL SI1 CI	PL SI3 CI
76	No Damage	0	0	0	0
248	No Damage	4	0	0	0
1046	No Damage	0	0	0	0
750		0	0	0	0
283	> 2 teeth macropitting	3	279	0	0
430	≥ 2 teeth macropitting	182	430	215	206
		PL RMS CI	PL FM4 CI	PL SI1 CI	PL SI3 CI
a	True Positive (TP)	185	709	215	206
a+c	Number of Faults	713	713	713	713
a/(a+c)	True Positive Rate(TPR)	0.26	0.99	0.30	0.29
b	False Positive (FP)	4	0	0	0
b+d	Number of healthy	1370	1370	1370	1370
b/(b+d)	False Positive Rate (FPR)	0.00	0.00	0.00	0.00
d	True Negative (TN)	1366	1370	1370	1370
a+d		1551	2079	1585	1576
a+b+c+d	Total Number of cases	2083	2083	2083	2083
(a+d)/(a+b+c+d)	Accuracy	0.74	1.00	0.76	0.76

Table 9 lists the calculated values for each parameter. The red arrows shown where the data came from to make the calculations for Pinion Left RMS. Reviewing the entire table, the only CI that achieved a detection rate of >90 percent at a threshold of 4.5 was FM4. However, the false alarm rate was 0 percent for all four CIs. That is why the accuracy is >70 percent for each CI. Although the detection rate was poor, the false alarm rate was good. The detection rate and false alarm rate are equally weighted for these calculations.

Note the importance of knowing the damage state to reliably assess the damage detection rate and false alarm rate. This is a huge challenge when assessing the helicopter data since the transition region from no damage to damage state is unknown. Only data before and after replacement is available. How the before replacement data is handled can have a significant effect on the detection rate. For example, if all of the data prior to replacement is assumed to be faulted, but the fault occurred a few acquisitions before replacement, the number of acquisitions prior to replacement can decrease the true positive rate of the CI. If it is assumed that no damage occurred in the timeframe after replacement, but damage occurred a 2nd time causing CI values to increase, the false alarm rate will also increase. Due to the unknown state of the gear sets, this analysis will not be presented for the helicopter data because the results will be skewed based on the gear state assumptions made and the number of acquisitions.

Table 10 lists the TPR, FPR and accuracy of condition indicators RMS, FM4, SI1 and SI3 for the eight rig tests. FPR cells highlighted in light green indicate a false alarm rate that is less than 0.10 or 10 percent. The false alarm rate was below this limit for all tests except test L1818R1616. TPR cells and accuracy cells highlighted in dark green indicate a CI damage detection rate and accuracy rate of greater than 0.9 or 90 percent. TPR cells and accuracy cells highlighted in blue indicate a CI damage detection rate and accuracy rate of greater than 75 percent but less than 90 percent. Reviewing Table 10, the CI with the highest detection rate and lowest false alarm rates across all tests was pinion SI1.

TABLE 10.—TPR, FPR AND ACCURACY FOR RIG TESTS

Test		PL RMS CI	PL FM4 CI	PL SI1 CI	PL SI3 CI	GL RMS CI	GL FM4 CI	GL SI1 CI	GL SI3 CI
L4545R5050	True Positive Rate(TPR)	0.26	0.99	0.3	0.29				
	False Positive Rate (FPR)	0	0	0	0				
	Accuracy	0.74	1	0.76	0.76				
L1515R5050	True Positive Rate(TPR)	0	0.43	0	0				
	False Positive Rate (FPR)	0	0	0	0				
	Accuracy	0.1	0.49	0.1	0.1				
L3030R5050	True Positive Rate(TPR)	0.35	0	0.87	0.83	0.05	0.76	0	0
	False Positive Rate (FPR)	0	0	0	0	0	0	0	0
	Accuracy	0.41	0.09	0.88	0.85	0.14	0.78	0.09	0.09
L3535R5050	True Positive Rate(TPR)	0.08	0	0.97	0.99	0.08	0.78	0.05	0
	False Positive Rate (FPR)	0	0	0	0	0	0	0	0
	Accuracy	0.72	0.69	0.99	0.99	0.72	0.93	0.71	0.69
L1818R1616	True Positive Rate(TPR)	0	0	1	1	0	0	0.03	0
	False Positive Rate (FPR)	0.18	0	0	0	0.18	0	0	0
	Accuracy	0.17	0.21	1	1	0.17	0.21	0.24	0.21
L2121R1616	True Positive Rate(TPR)	0.39	0	0	0	0.18	0	0	0
	False Positive Rate (FPR)	0	0	0	0	0	0	0	0
	Accuracy	0.43	0.07	0.07	0.07	0.24	0.07	0.07	0.07
L2020R5050	True Positive Rate(TPR)	0.97	0	0.86	0	0.97	0	0	0
	False Positive Rate (FPR)	0	0	0	0	0	0	0	0
	Accuracy	0.98	0.32	0.91	0.32	0.98	0.32	0.32	0.32
L4040R5050	True Positive Rate(TPR)	0.03	0	0	0	0.03	0	0	0
	False Positive Rate (FPR)	0	0	0	0	0	0	0	0
	Accuracy	0.2	0.17	0.17	0.17	0.2	0.17	0.17	0.17

Based on this investigation on this specific component in the two systems, CI performance varied for different CIs across tests and helicopters and between both systems. Some CIs did respond the same across both systems for comparable failure modes. This investigation illustrated the importance of understanding both systems before assessing CI performance. The results also identified additional measurements that can be made to quantify and scale the effect of the differences between both systems on CI response. The steps outlined in the Figure 17 process flow chart identify the steps required. Additional research is required to define experimental methods to quantify the effect on CI response in both the helicopter and the test rig. Variance of CI response between tests and across helicopters requires further studies. Sensing methods closer to the source may improve rig vibration response and should be investigated.

11.0 Summary

This report documented the analysis of condition indicator data collected on helicopters and in a test rig when damage occurred to spiral bevel gear sets. The purpose of this work was to demonstrate the process for validation of rotorcraft HUMS for maintenance credits using rig tests. Tests were performed on spiral bevel gear sets in the NASA Glenn Spiral Bevel Gear Fatigue Test Rig to simulate the fielded failures of spiral bevel gears installed in a helicopter. HUMS data from these helicopters with fielded failures was processed with the same techniques applied to spiral bevel rig test data. The performance of four gear condition indicators were evaluated, RMS or DA1, FM4, SI3 or SI and SI1. CI data from eight rig tests and fourteen helicopters with gear and pinion damage were compared.

System factors that affect CI response were outlined and discussed that included system structural design. Transfer path measurements of both the test rig and the helicopter system identified structural characteristics unique to each system that affected CI response. These included the significantly lower amplitudes of frequency response functions measured on the test rig than the helicopter and the negative effect this test rig design had on vibration response.

A process flowchart was also provided that outlined the steps required for rig test CI validation. CI performance metrics of detection rates, false alarm rates and accuracy, were also calculated for the rig data using simple thresholds. Overall, SI1 performed the best at detecting spiral bevel gear set damage for both systems. However, CI response in the test rig and helicopter varied for different failure modes and was affected by system design and operating conditions. Understanding the dynamic response of the complete monitored system is critical if test rig data is to be used for helicopter HUMS validation. To do this, further research is required into better state awareness of the fielded systems, new sensing technologies responsive to failures and insensitive to environmental and operating conditions, experimental methods or models that quantify the effect of system design on CI response and new methods for setting thresholds that take into consideration the variance of each system.

Appendix A.—Acronyms

120KTA	120 knots true airspeed regime
AC	Advisory Circular
AGMA	American Gear Manufacturers Association
AH64	Apache helicopter
CAA	Civil Aviation Authority
CBM	Condition-Based Maintenance
CDF	Cumulative distribution function
CI	Condition Indicator
DA1	CI Diagnostic Algorithm 1 calculated from the root-mean-square
DOE	Design of Experiments
FAA	Federal Aviation Administration
FM4	CI Figure of Merit 4
FN	False Negative—damage present with no indication
FP	False Positive—no damage with indication
FPG101	Flat Pitch Ground 101 percent rotor speed regime
FPR	False Positive Rate or false alarm rate
FRF	Frequency Response Function
GL	Gear Left
GR	Gear Right
HUMS	Heath and Usage Monitoring System
KS	Kolmogorov–Smirnov statistical test
MDSS	Mechanical Diagnostic System Software
MSPU	Modern Signal Processing Unit
NASA	National Aeronautics and Space Administration
NGB	Nose Gearbox
OEM	Original Equipment Manufacturer
PC-GBS	Personal Computer Ground Based Station
PL	Pinion Left
PR	Pinion Right
RMS	CI Root-Mean-Square
SAA	Space Act Agreement
SI	CI Sideband Index calculated from averaged ± 3 sidebands around mesh
SI1	CI Sideband Index calculated from averaged ± 1 sidebands around mesh
SI3	CI Sideband Index calculated from averaged ± 3 sidebands around mesh
TDA	Tear Down Analysis
TN	True Negative—no damage and no indication
TP	True Positive—damage present and indicated
TPR	True Positive Rate or damage detection rate
TSA	Time Synchronous Averaged Data
U.S.	United States
UK	United Kingdom

References

1. Dempsey, Paula J.: Investigation of Spiral Bevel Gear Condition Indicator Validation via AC–29–2C Using Test Rig Damage Progression Test Data. NASA/TM-2014-218384, September 2014.
2. Dempsey, Paula J.; Wade, Daniel R.; Antolick, Lance J.; and Thomas, Josiah: Investigation of Spiral Bevel Gear Condition Indicator Validation via AC–29–2C Using Fielded Rotorcraft HUMS Data. NASA/TM-2014-218406, November 2014.
3. Federal Aviation Administration (FAA): FAA Advisory Circular (AC) 29–2C, Section MG–15, Airworthiness Approval of Rotorcraft (RC) Health Usage Monitoring Systems (HUMS). PS-ASW100-1999-00063, July 15, 1999.
4. Antolick, Lance J.; Wade, Daniel R.; and Brower, Nathan G.: Application of Advanced Vibration Techniques for Enhancing Bearing Diagnostics on a HUMS-Equipped Fleet. Proceedings of the Specialists' Meeting on Airworthiness, Condition Based Maintenance (CBM) and Health Usage Monitoring Systems (HUMS) sponsored by American Helicopter Society (AHS), Huntsville, Alabama, 11-13 February 2013.
5. Civil Aviation Authority: CAP 753, "Helicopter Vibration Health Monitoring (VHM) Guidance Material for Operators Utilizing VHM in Rotor and Rotor Drive Systems of Helicopters," First Edition June 2006. First Edition incorporating amendment 2012/01, August 2012.
6. Delgado, Irebert; Dempsey, Paula; Antolick, Lance; and Wade, Daniel: Continued Evaluation of Gear Condition Indicator Performance on Rotorcraft Fleet. Proceedings of the Specialists' Meeting on Airworthiness, Condition Based Maintenance (CBM) and Health Usage Monitoring Systems (HUMS) sponsored by American Helicopter Society (AHS), Huntsville, Alabama, 11-13 February 2013.
7. Antolick, Lance; Branning, Jeremy; Dempsey, Paula; and Wade, Daniel: Evaluation of Gear Condition Indicator Performance on Rotorcraft Fleet. Proceedings of the American Helicopter Society International 66th Annual Forum, Phoenix, Arizona, 11-13 May 2010.
8. Dempsey, Paula J.; and Brandon, E. Bruce: Validation of Helicopter Gear Condition Indicators Using Seeded Fault Tests. Proceedings of MFPT 2013 Conference, 13-17 May 2013, Cleveland, Ohio. NASA TM-2013-217872.
9. Handschuh, R.F.: Thermal Behavior of Spiral Bevel Gears. NASA TM-106518, 1995.
10. Handschuh, R.F.: Testing of Face-Milled Spiral Bevel Gears at High-Speed and Load. NASA/TM—2001-210743, 2001.
11. American Gear Manufacturers Association (AGMA): Bevel Gear Rating Suite by AGMA, Version 1.0.12, 2007.
12. American Gear Manufacturers Association: Appearance of Gear Teeth – Terminology of Wear and Failure. ANSI/AGMA 1010-E95.
13. Davies, D.P.; Jenkins, S.L; and Belben, F.R.: Survey of fatigue failures in helicopter components and some lessons learnt. *Engineering Failure Analysis 32* (2013) 134–151.
14. Stewart, R.M: Some useful data analysis techniques for gearbox diagnostics. Machine Health Monitoring Group, Institute of Sound and Vibration Research, University of Southampton, Report MHM/R/10/77, July 1977.
15. Zakrajsek, James J.: An investigation of gear mesh failure prediction techniques. NASA TM–102340, AVSCOM TM 89–C-005, 1989.
16. Zakrajsek, James J.; Handschuh, Robert F.; and Decker, Harry J: Application of fault detection techniques to spiral bevel gear fatigue data. Proceedings of the 48th Meeting of the Mechanical Failures Prevention Group. Office of Naval Research, Arlington, Virginia, pp. 93–104, 1994.

17. Wade, Daniel; Larsen, Chris; Roth, Richard; and Shelley, Stuart: Final Report of Bearing Condition Indicator Refinement Project, Technical Report Rdmr Ae-06-01, June 2011.
18. Massey, F. J.: The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*. Vol. 46, No. 253, pp. 68–78, 1951.
19. Islam, Anwarul; Dempsey, Paula; Feldman, Jason; and Larsen, Chris: Characterization and Comparison of Vibration Transfer Paths in a Helicopter Gearbox and a Fixture Mounted Gearbox. NASA/TM 2013-216586, September 2013.
20. Dempsey, Paula; Islam, Anwarul; Feldman, Jason; and Larsen, Chris: Investigation of Gearbox Vibration Transmission Paths on Gear Condition Indicator Performance. NASA TM—2013-216617, December 2013.
21. U.S. Army Research, Development & Engineering Command: Aeronautical Design Standard Handbook for Condition Based Maintenance Systems for U.S. Army Aircraft Systems. ADS-79B-HDBK, March 7, 2013.
22. Gareth, James; Witten, Daniela; Hastie, Trevor; and Tibshirani, Robert: An Introduction to Statistical Learning: With Applications in R. New York: Springer, 2013.